4-2003

# Genes Duplicated by Polyploidy Show Unequal Contributions to the Transcriptome and Organ-Specific Reciprocal Silencing

Keith L. Adams
Iowa State University

Richard Clark Cronn
United States Department of Agriculture

Ryan J. Percifield
Iowa State University

Jonathan F. Wendel
Iowa State University, jfw@iastate.edu

# Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing

Keith L. Adams*, Richard Cronn†, Ryan Percifield*, and Jonathan F. Wendel*‡

*Department of Botany, Iowa State University, Ames, IA 50011; and †U.S. Department of Agriculture Forest Service, Pacific Northwest Research Station, Corvallis, OR 97331

Communicated by John F. Doebley, University of Wisconsin, Madison, WI, January 31, 2003 (received for review November 20, 2002)

Most eukaryotes have genomes that exhibit high levels of gene redundancy, much of which seems to have arisen from one or more cycles of genome doubling. Polyploidy has been particularly prominent during flowering plant evolution, yielding duplicated genes (homoeologs) whose expression may be retained or lost either as an immediate consequence of polyploidization or on an evolutionary timescale. Expression of 40 homoeologous gene pairs was assayed by cDNA-single-stranded conformation polymorphism in natural (1- to 2-million-yr-old) and synthetic tetraploid cotton (*Gossypium*) to determine whether homoeologous gene pairs are expressed at equal levels after polyploid formation. Silencing or unequal expression of one homoeolog was documented for 10 of 40 genes examined in ovules of *Gossypium hirsutum*. Assays of homoeolog expression in 10 organs revealed variable expression levels and silencing, depending on the gene and organ examined. Remarkably, silencing and biased expression of some gene pairs are reciprocal and developmentally regulated, with one homoeolog showing silencing in some organs and the other being silenced in other organs, suggesting rapid subfunctionalization. Duplicate gene expression was examined in additional natural polyploids to characterize the pace at which expression alteration evolves. Analysis of a synthetic tetraploid revealed homoeolog expression and silencing patterns that sometimes mirrored those of the natural tetraploid. Both long-term and immediate responses to polyploidization were implicated. Data suggest that some silencing events are epigenetically induced during the allopolyploidization process.

Nearly all higher eukaryotes have genomes that exhibit extensive gene redundancy. Much of this redundancy seems to have arisen from genome doubling, or polyploidy. Polyploidy has been particularly significant in the evolutionary history of vertebrates (1–3), yeast (4), and flowering plants (5–7). Accordingly, most of these genomes contain duplicated chromosomes or chromosomal segments that reflect ancient or recent rounds of polyploidy. Polyploidy is especially prevalent in plants, where at least 50% and perhaps up to 95% of angiosperms have experienced one or more episodes of chromosome doubling in their evolutionary history (8, 9). Polyploidy is widely perceived to provide raw material for the origin of physiological and morphological novelty (10, 11).

Investigations of allopolyploid plants (formed from two different diploids) have shown that the merger of two genomes in a common nucleus may be accompanied by considerable genomic reorganization (reviewed in refs. 5, 12, and 13). Rapid and nonrandom changes, including sequence elimination, have been inferred in wheat and *Brassica* polyploids (14–18), and these changes may be accompanied by methylation alterations (18).

Cotton (*Gossypium*) has been developed as a particularly useful group for studies of polyploidy (19). Polyploidization between an A-genome diploid and a D-genome diploid ≈1.5 million years ago (20, 21) created an AD tetraploid lineage that has since diversified into five species, including the commercially important *Gossypium hirsutum* (upland cotton) and *Gossypium barbadense* (Pima cotton) (see Fig. 1). Previous studies have demonstrated the possibility of intergenomic "cross-talk" after polyploid formation. For example, transposable elements and other repeated sequences from one genome have colonized the alternative genome subsequent to polyploidization (22), whereas other repeated sequences experienced interlocus homogenization, whereby sequences from one genome were overwritten by homoeologous sequences from the other genome (23). The above studies from *Gossypium* and other model plant systems highlight the dynamic nature of polyploid genomes and the relatively poorly understood and some times non-Mendelian mechanisms that often characterize gene and genome evolution in polyploids (5, 13, 24).

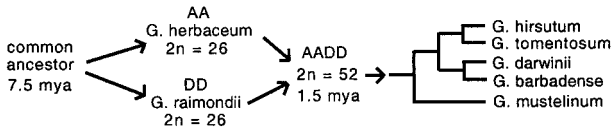

Fig. 1. Schematic representation of the phylogenetic history of diploid and allopolyploid *Gossypium*. Polyploid formation occurred ≈1.5 million years ago after hybridization between A genome and D genome diploids (19). After formation, the ancestral polyploid radiated into lineages represented now by the five species shown.

Studies of gene expression in natural and synthetic plant polyploids have shown that some genes are silenced after polyploidization, as shown first for isozyme loci (25–27). Ribosomal RNA arrays from one parent may be silenced in some organs (28), although both parental rRNA sets are expressed in floral organs of *Brassica napus* (29). Studies of several hundred loci in *Arabidopsis suecica* and wheat using cDNA-amplified fragment-length polymorphism screens have documented silencing of protein coding genes (30–32). In most cases one homoeolog was silenced, although both homoeologs of some genes were silenced in wheat.

To date, relatively little information exists regarding the proportional contributions of two newly merged genomes to the transcriptome of allopolyploids, both overall and on a gene-by-gene basis. In part, this absence of evidence reflects experimental difficulties inherent in distinguishing transcripts from two, usually similar parental genomes. Expression of homoeologous gene pairs has not been compared in different organ types or between natural and synthetic genotypes. Here, we survey transcript accumulation for 40 pairs of genes duplicated by polyploidy (homoeologs) in cotton. Our results show that, although many homoeologs contribute approximately equally to the transcriptome, a surprisingly high percentage of genes exhibit silencing or biased expression that is developmentally regulated, both in natural and synthetic polyploids. We show that in a few cases alternative homoeologs have been reciprocally silenced in different organs, suggesting subfunctionalization (33, 34).

Abbreviations: SSCP, single-stranded conformation polymorphism; RT, reverse transcription; dpa, days postanthesis.

See commentary on page 4369.

‡To whom correspondence should be addressed. E-mail: jfw@iastate.edu.

Materials and Methods

Plant Materials and Nucleic Acid Extractions. Natural allotetraploids included *G. hirsutum* cultivar TM1, *Gossypium mustelinum* (accession no. AD4-15C), and *Gossypium darwinii* (accession no. AD5-PW45). Models of their progenitor diploids (19) included the A-genome *Gossypium herbaceum* (accession no. A1-73) and the D-genome *Gossypium raimondii* (Wendel laboratory stock). A synthetic allotetraploid 2(A_2D_1) was also studied, this generated by colchicine-doubling a sterile diploid hybrid formed between the A-genome species *Gossypium arboreum* and the D-genome species *Gossypium thurberi* (35). Because exact parental plants of this synthetic were not available, *G. arboreum* (cv. AKA8401) and *G. thurberi* (accession nos. 8 and 5) were used as models of the parents. All plants were grown in a greenhouse under common conditions.

For RNA isolation, the following vegetative organs were collected: young leaves; cotyledons from seedlings 7 days after first appearance; stems from seedlings after appearance of the first true leaf; and roots from a mature plant. Floral organs, collected from multiple plants on the day of anthesis (flower opening) between 9:30 a.m. and noon on several days, included bracts (epicalyx), calyx, petals, whole stamens, and stigma + style (referred to in figures as ''carpels''). Ovules were collected at 5, 10, 15, and 20 days postanthesis (dpa) from multiple plants.

DNAs were extracted by using the Qiagen (Valencia, CA) DNeasy kit. RNAs were extracted as described (36) with a few modifications (see *Supporting Text*, which is published as supporting information on the PNAS web site, www.pnas.org.). RNA concentrations were estimated by using a spectrophotometer. Expression assays in ovules used mixed RNAs from 5, 10, 15, and 20 days postanthesis (dpa) in similar amounts. RNAs were treated with Dnase I before reverse transcription using the DNA-free kit (Ambion, Austin, TX).

Gene Amplification. The 40 genes selected (Table 1) were from previous molecular phylogenetic studies and ongoing investigations (21, 37–39). Sequence and phylogenetic analysis confirmed homoeology among the duplicated copies (21). Reverse transcription (RT) used 2 μg of RNA and was performed with the RETROscript kit (Ambion) according to the manufacturer's instructions. As controls for DNA contamination, reactions were also performed without RT (RT^-), side-by-side with experimental reactions. All reactions were followed by treatment with Rnase A for 20 min at 37°C. One-twentieth of the cDNAs created by first strand synthesis were used in PCR reactions with 0.5 μM each primer, 2.5 mM $MgCl_2$, and *Taq*DNA polymerase. Reaction volumes were 30 μl, and cycling was done in a MJ PTC-100 thermocycler for 2 min at 94°C, followed by 30 cycles of 30 s at 94°C, 30 s at 52–59°C, 1 min at 72°C, and then a final 6 min extension at 72°C. Primers (Table 4, which is published as supporting information on the PNAS web site) were designed to match the sequences of both duplicates for each gene from *G. hirsutum*. *Adh* gene primers were designed to specifically amplify each gene family member separately.

Single-Stranded Conformation Polymorphism (SSCP) Analysis. SSCP analysis was done as in ref. 40. Urea concentrations ranged from 2–10%, and electrophoresis was done either at 4°C or room temperature, as empirically determined for resolution of each of the 40 gene pairs studied (Table 6, which is published as supporting information on the PNAS web site). Most SSCP reactions were conducted more than once. Band quantification was accomplished by using a Molecular Dynamics Storm 840 PhosphorImager and IMAGEQUANT software. To test the reproducibility of the cDNA-SSCP assay, replicates with *G. hirsutum* were performed from the RT stage forward with genes *adhA*, *adhD*, and A1520. For most inferred silenced genes, direct sequencing of RT-PCR products (using an ABI 377 DNA sequencer) was performed for confirmation.

Table 1. cDNA-SSCP analysis of homoeolog expression in ovules

Gene	Putative function/function	
One homoeolog silenced		
adhE	Alcohol dehydrogenase E	At only
Strongly biased expression (85/15% and greater)		
adhD	Alcohol dehydrogenase D	At > Dt
G8	Flavonoid 3′ hydroxylase	Dt > At
B5	Oxalate oxidase	At > Dt
adhA	Alcohol dehydrogenase A	Dt > At
Biased expression (60/40% to 84/26%)		
D5	Sugar transporter	At > Dt
D7	Root hair defective 3 homolog	At > Dt
E6	Potassium transport protein	Dt > At
H12	Auxin-alanine hydrolase	At > Dt
A1520	Unknown	Dt > At
Equivalent expression 50/50 to 59/41%		
A6	Unknown	
A1751	Subtilisin-like protease	
B2	Kinesin heavy chain	
B3	Enolase	
B7	Polyubiquitin	
B8	CAAX-prenyl protease	
C1	Vacuolar ATP synthase	
C3	Syntaxin	
C7	Pollen allergen-like protein	
CesA2	Cellulose synthase	
D1	LIM-domain transcription factor	
E1	Flavonoid 3′,5′-hydroxylase	
E5	Pollen surface protein	
E9	Protein phosphatase	
E11	β-D-glucan exohydrolase	
F4	Ethylene receptor	
F10	Xyloglucan endotransglycosylase	
F12	COP-1 interacting protein	
G3	Quinone oxidoreductase	
G6	CCAAT-binding transcription factor	
G7	Dehydration-responsive protein	
G11	Sulfate transporter	
G1134	Unknown	
G1262	P-glycoprotein	
H5	Gibberellin-regulated protein	
H6	Auxin repressed protein	
myb1	myb R2R3 transcription factor 1	
myb2	myb R2R3 transcription factor 2	
myb3	myb R2R3 transcription factor 3	
Rac13	GTP-binding protein	

GenBank numbers are provided in Table 5, which is published as supporting information on the PNAS web site.

Results

Transcript Contributions of the Two Genomes in Ovules of Allotetraploid *G. hirsutum*. The proportional contribution to the transcriptome of transcripts derived from both copies of 40 homoeologous gene pairs (designated At and Dt) was examined to determine whether there is a bias in transcript levels from one of the two genomes and whether alterations in expression have occurred postpolyploidization. In the initial screen, transcript levels were assayed by using cDNA-SSCP (40) in whole ovules (including attached fibers) of cultivated cotton (*G. hirsutum*). RT-PCR was performed on cDNA templates from the natural allotetraploid *G. hirsutum* and model progenitor diploids. PCR was performed on genomic DNA from *G. hirsutum* by using the same primers to locate expected positions of RT-PCR products on SSCP gels, and to serve as controls for amplification bias. All cDNA templates were

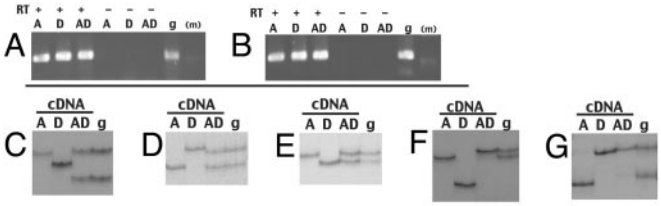

Fig. 2. Homoeologous gene pair expression in cotton ovules. ''A'' and ''D'' indicate diploid A- and D-genome species *G. herbaceum* and *G. raimondii*, respectively; ''AD'' indicates allotetraploid *G. hirsutum;* and ''g'' indicates genomic DNA control from *G. hirsutum*. (*A* and *B*) Examples of RT-PCR results. Reactions were performed with (+) or without (−) RT. Genes include *adhA* (*A*) and gene F12 (*B*); (*C–G*) cDNA-SSCP analysis. Genes include (*C*) COP-1 interacting protein (F12); (*D*) syntaxin (C3); (*E*) sulfate transporter (G11); (*F*) alcohol dehydrogenase A; and (*G*) flavonoid 3′ hydroxylase (G8). Note equal expression of homoeologs in *C–E*, but biased expression (unequal intensities of the two bands in the AD lane) in *F* and *G*. In some cases, the tetraploid sequences migrate differently from the corresponding diploid sequences because of mutations that have occurred postpolyploidization or because the diploid species are the best living models rather than the actual progenitors of the natural allopolyploid.

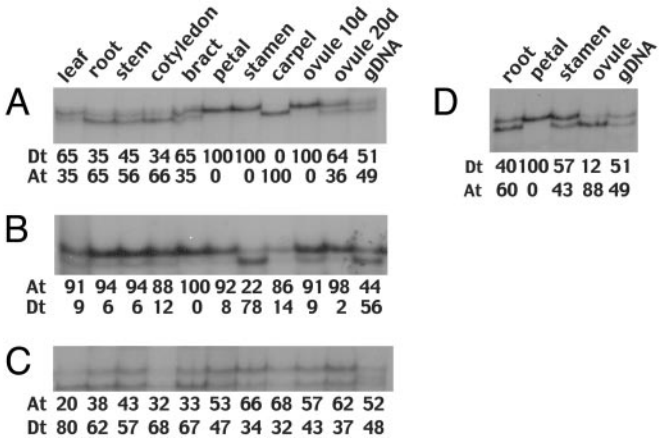

Fig. 3. Transcript levels from duplicated gene pairs in multiple organs of *G. hirsutum*, as determined by cDNA-SSCP analysis. ''Dt'' and ''At'' denote homoeologous genes from the A or D genome of the natural allopolyploid *G. hirsutum*, for which the numbers indicate transcript percentages as determined by PhosphorImager. (Note that for some genes At ran further on the gels and for other genes Dt ran further.) Genes include (*A*) *adhA* alcohol dehydrogenase A; (*B*) *adhD* alcohol dehydrogenase D; (*C*) putative protein A1520; and (*D*) oxalate oxidase (gene B5). One strand is shown for RT-PCR products that were labeled on both strands.

checked for contaminating genomic DNA by using RT^- controls containing all RT reagents except reverse transcriptase. No DNA contamination was detected (Fig. 2).

The proportion of RT-PCR products derived from each copy in a homoeologous gene pair was assayed by SSCP gels and PhosphorImager quantification. This method has been shown to yield quantitative estimates of transcript ratios in template pools ranging from equal amounts to ≈100:1 (40). SSCP gel conditions were optimized for each gene by adjusting running temperature and urea concentration so that homoeologs would resolve and single stable conformations would form. Fig. 2 shows examples of SSCP gels for genes that exhibit either approximately equal expression of both homoeologs (Fig. 2 *C–E*) or an appreciable expression bias toward one homoeolog (Fig. 2 *F* and *G*). Transcript accumulation was considered to be approximately equal if transcript amounts from the two homoeologs ranged from 50/50 to 59/41, whereas an expression bias was inferred when the transcript ratio for the two homoeologs was 60/40 or greater (Table 1; 60/40 was an arbitrary cut-off, selected to ignore slight departures from equal expression). By using this scoring system, nine genes were interpreted to show biased expression, five toward the At homoeolog and four toward the Dt homoeolog. The Dt copy of one gene (*adhE*) was silenced in ovules. Of the 10 genes that show biased expression or silencing, 6 are from the A genome and 4 are from the D genome. Thus, there does not seem to be preferential expression of genes from one of the two genomes in *G. hirsutum* ovules. Six of the 10 genes are enzymes, two are transporters, and two have other functions.

Reciprocal, Organ-Specific Expression and Silencing of Homoeologs in *G. hirsutum*. To explore possible organ-specific partitioning of homoeologous gene expression in *G. hirsutum*, 10 organs were selected for further study. We were especially interested in determining whether there has been reciprocal silencing of homoeologs from the At and Dt genomes in different organs such that both genes remain functional in different parts of the plant, suggestive of partitioning of ancestral function. Transcript levels for 16 gene pairs were examined in 8–10 organs by cDNA-SSCP analysis, and transcript levels for two additional genes were examined in 4 organs. Eleven of the 18 genes showed homoeolog silencing or biased expression in at least 1 organ type. Perhaps the most striking example is *adhA*, where proportional transcript abundance from the two homoeologs varied from nearly equal to exclusively from one duplicate or the other. As shown in Fig. 3*A*, transcripts from both genomes were detected in vegetative organs (although mostly in a biased fashion; note preferential expression of the Dt homoeolog in leaves and bracts and bias toward At expression in cotyledons and roots), whereas in some floral organs, there has been complete reciprocal silencing in different floral whorls. In petals and stamens, the At homoeolog has been silenced, but the reverse is true in carpels, where no transcripts were detected from the Dt homoeolog.

Other genes, including *adhD* (Fig. 3*B*) and A1520 (Fig. 3*C*), also showed biased expression toward the At homoeolog in some organs and the Dt homoeolog in other organs. This pattern was particularly striking for *adhD*, which showed mostly expression of At in all organs assayed except stamens, where over three-fourths of the transcripts were derived from the alternative gene copy. A1520 showed preferential expression of Dt in all vegetative organs examined, but preferential expression of At in reproductive organs and ovules. Gene B5 (oxalate oxidase; assayed only in roots, petals, stamens, and ovules), showed silencing of At in petals but strong preferential expression of At in ovules (Fig. 3*D*). Gene G7 (a dehydration responsive protein) was strongly biased toward the At homoeolog in petals, but toward Dt in carpels (Table 2). In contrast, *adhE* showed biased accumulation of Dt transcripts in all organs examined, except that we were unable to detect any transcripts from either homoeolog in stamens and carpels (Table 2).

Some genes, e.g., G8 (flavonoid hydroxylase), displayed a bias toward expression of the same homoeolog in some organs and equal expression in others (Table 2). Strongly biased expression only in a single organ was detected for some genes. Gene D7 (root hair defective 3 homolog) displayed an expression bias only in roots, whereas gene G1262 (a *p*-glycoprotein) expression was biased only in stems (Table 2). Transcripts from gene G1262 were detected only in roots, stems, and stamens; this gene could have a restricted expression pattern in the diploid progenitors. Finally, seven genes showed equivalent levels of homoeolog expression in all tested organs. These genes included A6, B2, B8, E5, F12, G6, and G1134 (Table 1).

To assess biological variation (variation between plants and environmental fluctuations) in homoeolog expression, RNA was extracted from other *G. hirsutum* individuals at a different time, and transcript levels were assayed in floral organs for the genes *adhA* and *adhD*, which showed particularly interesting expression patterns. Expression ratios of *adhD* homoeologs were equivalent to those of the first survey. Similarly, the *adhA* gene silencing observed

Table 2. cDNA-SSCP analysis of homoeologous gene pair expression in 10 organs of *G. hirsutum*

	Leaf	Root	Stem	Cot.	Bract	Petal	Stamen	Carpel	10d ov.	20d ov.	gDNA
Gene	At, Dt	At, Dt	At, Dt	At, Dt	At, Dt	At, Dt	At, Dt	At, Dt	At, Dt	At, Dt	At, Dt
adhA (1)	35, 65	65, 35	56, 44	66, 34	35, 65	0, 100	0, 100	100, 0	0, 100	64, 36	51, 49
adhA (2)	35, 65	62, 38	64, 36	71, 29	42, 58	0, 100	0, 100	100, 0	0, 100	67, 33	50, 50
adhD (1)	94, 6	92, 8	80, 11	91, 9	100, 0	97, 3	19, 81	84, 16	87, 13	89, 11	41, 59
adhD (2)	91, 9	94, 6	94, 6	88, 12	100, 0	92, 8	22, 78	86, 14	91, 9	98, 2	44, 56
A1520 (1)	20, 80	38, 62	43, 57	32, 68	33, 67	53, 47	66, 34	68, 32	57, 43	62, 38	52, 48
A1520 (2)	20, 80	—	47, 53	—	38, 62	53, 47	63, 37	62, 38	52, 48	—	50, 50
adhE	7, 93	4, 96	6, 94	7, 93	22, 78	9, 91	—	—	0, 100	0, 100	—
myb1	62, 38	44, 56	42, 58	82, 18	66, 34	66, 34	55, 45	62, 38	46, 54	42, 58	50, 50
G8	11, 89	13, 87	29, 71	24, 76	9, 91	32, 68	48, 52	43, 57	24, 76	46, 54	50, 50
D5	66, 34	72, 28	63, 37	71, 29	71, 29	—	—	—	59, 41	59, 41	53, 47
D7	51, 49	23, 77	39, 61	49, 51	53, 47	46, 54	41, 59	55, 45	59, 41	47, 53	48, 52
G7	70, 30	61, 39	—	42, 58	69, 31	75, 25	—	14, 86	32, 68	51, 49	—
B5		60, 40				0, 100	43, 57				49, 51
G1262		61, 39	87, 13								56, 44

At and Dt values represent the percentage of transcripts derived from the At genome and Dt genome, respectively, of allotetraploid *G. hirsutum.* Cot, cotyledon; ov, ovule. Dashes indicate expression ratio not determined. Numbers in parentheses indicate replicates. Gene names correspond with Table 1.

in petals, stamens, and carpels was also evident in the replicates. However, the bias against At transcript accumulation in bracts was even more severe than in the original survey, showing that quantitative estimates of transcript ratios may be sensitive to environmental and/or biological factors.

To evaluate the possibility that gene silencing observed in the polyploid actually originated in the diploid progenitors and has merely been retained since polyploid formation 1–2 million years ago, we assayed transcript levels in homologous organs from model plants representing the A genome and D genome diploid progenitors (described above.) Transcripts were readily detected in all tested organs (Fig. 6, which is published as supporting information on the PNAS web site). Most importantly, the *adhA* gene silencing scored in allopolyploid cotton is inferred to have arisen during or after polyploid formation. Cases of biased expression without full silencing also probably reflect alterations in expression levels postpolyploidization, although in some cases it is possible that diploids varied in expression levels and that this quantitative variation was maintained in descendant polyploids.

Homoeolog Expression and Silencing in a Synthetic Allotetraploid. To address whether organ-specific alterations in homoeolog expression occur during or soon after polyploid formation or whether they arise more slowly on an evolutionary timescale, we studied homoeolog expression in a synthetic allotetraploid of similar genomic composition to the natural polyploid cottons. Expression of four homoeologous gene pairs was assayed by cDNA-SSCP in five floral organs of two plants. Replicates were performed by using new RNA extractions from tissue collected at different times from one of the two plants used originally, and genes *adhA* and G8.

Patterns for *adhA* in the synthetic allopolyploid were remarkably similar to those of *G. hirsutum* in stamens and carpels: most transcripts in the stamens were derived from the Dt homoeolog, and almost all transcripts in carpels were derived from the At homoeolog (Fig. 4*A*). Thus, reciprocal expression biases can occur during or soon after polyploid formation. Gene G8 also showed organ-specific reciprocal silencing in the synthetic: nearly all transcripts in bracts were from the Dt homoeolog whereas there was near-exclusive At bias in petals and stamens (Fig. 4*B*; the faint pair of extraneous bands in the stamen lane was not reproducible.) G8 expression in the synthetic allopolyploid contrasts with the natural allopolyploid, where there was bias toward At or near-equal expression in all organs. Expression patterns for *myb1* and G7 (Table 3) were similar, in that Dt was silenced in petals, in contrast to *G. hirsutum*.

To evaluate the possibility that the expression alterations observed in the synthetic polyploid actually reflect inheritance of divergent expression patterns of diploid progenitors, we assayed transcript levels in homologous organs from the diploids. In all cases (*adhA*, G8, *myb1*), transcripts were detected by RT-PCR in the A and D diploids (Fig. 7, which is published as supporting information on the PNAS web site). Thus, the expression alterations scored in the synthetic allopolyploid are inferred to have arisen during or soon after allopolyploid formation.

Expression Alterations During Evolution of *Gossypium* Polyploids. If gene expression alteration is primarily a consequence of genomic merger and/or if gene expression changes arose near the time of polyploid formation ≈1.5 million years ago before the diversification of the five natural polyploid species, one might expect the different wild allopolyploids to display expression profiles similar to those observed for *G. hirsutum*. Alternatively, if expression alteration arises more erratically and on an evolutionary timescale, one would expect to observe a more random pattern of silencing and biased expression among the various allopolyploids. To address this issue, we assayed *adhA* and *adhD* expression in four organs of two species (*G. mustelinum* and *G. darwinii*) that represent the other two branches of the phylogenetic tree of *Gossypium* polyploids (Fig. 1; ref. 19).

AdhA expression in carpels and floral bracts was identical in all three species (Fig. 5). Because the Dt homoeolog has been silenced in carpels in all three species, we infer that silencing originated in the common ancestor of the polyploids, perhaps during or soon after polyploidization. In contrast to the shared interspecific silencing observed in carpels, homoeolog expression in stamens and petals differed among the three natural polyploids. For example, silencing of the At homoeolog in stamens was observed only in *G. hirsutum*. Expression for *adhD* was comparable in all three natural

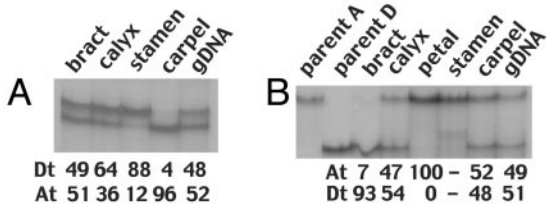

Fig. 4. Transcript levels of homoeologous gene pairs in synthetic allopolyploid cotton. ''Dt'' and ''At'' indicate homoeologous genes from the A or D genome. Numbers indicate transcript percentages as determined by PhosphorImager. Dashes indicate lanes that were not scored on the gel shown. (*A*) *adhA*, alcohol dehydrogenase; (*B*) flavonoid 3′ hydroxylase (gene G8).

Table 3. cDNA-SSCP analysis of homoeolog expression in a synthetic *Gossypium* allotetraploid

	Bract	Calyx	Petal	Stamen	Carpel	gDNA
Gene	At Dt	At, Dt	At, Dt	At, Dt	At, Dt	At, Dt
adhA (1)	51, 49	37, 63	—	12, 88	96, 4	48, 52
adhA (2)	53, 47	39, 61	—	9, 91	98, 2	—
G8 (1)	7, 93	47, 53	100, 0	6, 94	52, 48	49, 51
G8 (2)	5, 95	—	100, 0	4, 96	—	—
myb1	52, 48	52, 48	100, 0	—	56, 44	45, 55
G7	49, 51	74, 26	100, 0	65, 35	58, 42	—

At and Dt represent the percentage of transcripts derived from the At and Dt genomes, respectively, of the allotetraploid. Numbers in parentheses are replicates. Dashes indicate expression ratio not determined.

polyploids, although proportional contributions were variable, particularly in the carpels (Fig. 5).

Discussion

Contributions of Duplicated Genomes to the Allopolyploid Transcriptome. Allopolyploid speciation entails the merger of two divergent genomes in a common nucleus, doubling the number of genes. In principle, both genomes could contribute equally to the transcriptome, or alternatively, there may be preferential transcription of one genome due to genomic differences that affect the transcription machinery. Here, we examined expression of 40 gene pairs in at least one organ of allopolyploid cotton. Of these, 13 showed biased expression or silencing, whereas approximately equal transcript amounts were detected for both homoeologs of the other 27 gene pairs. Of those genes exhibiting an expression bias, 5 showed bias toward At, 2 showed bias toward Dt, and 6 showed bias toward the At homoeolog in some organs and toward Dt in other organs. Thus, based on a sampling of 40 duplicated gene pairs, there does not seem to be preferential transcription of genes from one of the two genomes of *G. hirsutum*. Although this result may at first seem unremarkable, we note that no data have evaluated the proportional contribution of two genomes to the transcriptome of an allopolyploid.

A High Degree of Expression Alteration That Is Developmentally Regulated. Although there seems to be no global genomic bias in transcription, expression of individual genes varies greatly, and transcript levels of almost one-third of the individual gene pairs examined revealed appreciable bias toward one homoeolog or the other in at least one organ. Gene silencing and expression biases varied greatly among organs for 10 of the 18 genes in which expression was studied in multiple organs of *G. hirsutum*, and for all 4 genes studied in the synthetic allotetraploid. Six genes showed biased expression/silencing of homoeologous gene pairs that is reciprocal in different organs. These results show that the genomic response to polyploidy varies in different parts of the plant, varies widely among genes, and can be developmentally regulated in heretofore unanticipated ways. The most spectacular examples of the latter are genes that show developmentally regulated reciprocal silencing of alternative homoeologs, where there is minimal to no transcription of one member of a duplicated gene pair in some organs and a similar absence of transcription of its duplicate in other parts of the plant. This is an unprecedented observation, but we predict that, as more studies are conducted in a comparable manner, it will turn out to be a frequent consequence of genomic merger.

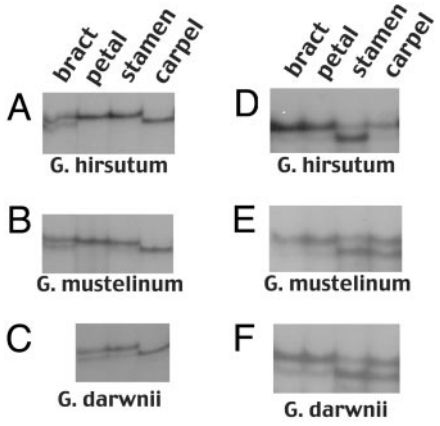

Fig. 5. Expression of homoeologous gene pairs in three species of allopolyploid cotton. Phylogenetic relationships among the species used are shown in Fig. 1. (*A–C*) *adhA*. (*D–F*) *adhD*.

Given the fact that most of the organs examined in our study consist of several tissues and many cell types, it may be that the degree of departure from equal expression is even more profound than we observed in our surveys of whole organs. For example, biased expression could reflect differences in transcript contributions among cell types (i.e., expression in some cells but silencing in others), or it could reflect the contributions of every cell in the organ. It will be of interest to explore this question further by using cell- and tissue-specific message pools. Expression of homoeologous genes may vary during the development of a particular organ. For example, *adhA* expression in ovules at 10 dpa was significantly different from at 20 dpa (Fig. 3). Environmental factors (such as light, photoperiod, temperature, and stress) may play a role in expression of homoeologous gene pairs; we are beginning to explore this interesting possibility.

Mechanisms of Developmentally Regulated Expression Variation and Silencing. The observation of organ-specific expression alteration in synthetic allopolyploid cotton implicates one or more epigenetic mechanisms as the cause of the observed changes. The formal alternative of polyploidy-induced mutational modification is considered improbable because few such changes are observed in synthetic allopolyploid cottons (41) and because it is difficult to envision how genome rearrangements or nucleotide substitutions might spontaneously occur to yield such a high frequency of organ-specific alteration in transcript accumulation. Epigenetic changes in other plants have been shown to accompany polyploidy, and experiments using a DNA methyltransferase inhibitor implicated modifications in DNA cytosine methylation (31, 42). Related possible epigenetic causes include histone modifications and positional effects from higher-order changes in chromatin structure (43, 44). Such changes might be due either to myriad possible intergenomic interactions that influence the transcription machinery or perhaps to the physical requirements of housing two genomes in a single nucleus. These epigenetic factors are not mutually exclusive, and it is possible that mechanisms vary by gene and involve multiple underlying controls. Irrespective of cause, epigenetic mechanisms may account for expression changes in natural cotton polyploids as well. In this case, however, additional modifications may arise from a slower accumulation of genetic mutations in gene regulatory elements that abolish or reduce expression in particular organs.

The natural and synthetic cotton allopolyploids used in this study result from both a hybridization event that united two divergent genomes (differing in size by a factor of two) and a change in ploidy from diploid to tetraploid. Both phenomena may influence gene expression. For example, studies of a synthetic *Arabidopsis* polyploid (30, 42) have shown gene silencing in plants that resulted only from genome merger (i.e., both parents were tetraploid). Silencing has also been associated with ploidy changes (45). It will be interesting to determine, by using diploid hybrids, whether the organ-specific gene silencing documented in allopolyploid cottons is due to genome merger, ploidy change, or both.

Expression Changes Accompany the Onset of Polyploidy and Continue Over Evolutionary Time. Our results demonstrate that organ-specific expression changes commonly arise with the onset of polyploidization. Interestingly, all four genes examined in the synthetic allopolyploid also showed silencing and/or biased expression in natural *G. hirsutum*. For *adhA*, the expression changes are similar in the natural and synthetic polyploids, whereas for genes G8 and *myb1* they are somewhat different. Initial data for *adhA* in a second independently created AD-genome allotetraploid shows a pattern of reciprocal silencing in stamens and carpels similar to that of the synthetic $2(A_2D_1)$ examined in this study (unpublished data). Thus, it is possible that some genes may be repeatedly silenced in independent polyploidization events by a directed process modulated by specific attributes of merging genomes.

Comparison of homoeologous gene expression in natural polyploid species that are derived from the same polyploidization event can provide a glimpse at the evolutionary timing and dynamics of gene silencing and expression alteration. Examination of *adhA* expression in three natural cotton polyploids suggest that the At copy was silenced in carpels before species radiation, perhaps concurrent with or soon after polyploid formation. The congruent *adhA* expression patterns in the wild and synthetic allopolyploids suggest that silencing arose at the time of allopolyploidization and raise the tantalizing possibility that there has been stable maintenance of an epigenetic mutation over at least 1 million years, or since allopolyploid *Gossypium* first arose. Such long-term epigenetic mutation maintenance has been documented for the *cycloidea* gene that affects floral symmetry (46). Alternatively, *adhA* expression patterns in carpels might be explained by an initial epigenetic modification that was then made permanent by a genetic lesion sometime during the evolution of the allotetraploid cottons. In contrast to the expression pattern in carpels, *adhA* expression in stamens showed silencing of the At homoeolog only in *G. hirsutum*, suggesting recent silencing in this species.

Evolutionary and Functional Significance. Relatively little is known regarding the functional consequences and evolutionary importance of expression modification after genome doubling, although theory predicts that one of the consequences is a partitioning of aggregate ancestral function between the two duplicates (33, 34, 47). This process of subfunctionalization is conceived to be one that operates on an evolutionary timescale, requiring fixation of mutations in regulatory regions or functional domains such that complementary degenerate mutations arise. If the reciprocal and complete silencing of *adhA* homoeologs observed in *G. hirsutum* (Fig. 3*A*) represents subfunctionalization, then it is probably the most evolutionarily recent example reported.

The expression alterations that accompany polyploidization, many of which are likely to be epigenetic, may be both evolutionarily stable and latently variable. When extended to entire duplicated genomes, such epigenetic differential homoeolog silencing may create vastly increased reservoirs of physiological variation on which selection might act, in the process perhaps fixing advantageous epigenetic states by more slowly accumulating mutational means. This process may be evidenced in some of our data, where, for example, the developmentally regulated silencing of *adhA* may be construed as an organ-specific partitioning of function, which is both an immediate consequence of polyploidy formation and whose expression state has, in some cases, remained unaltered, perhaps by mutation, to the present.

From a functional perspective, why might genes encoding similar or even identical proteins be silenced? We note that sequence divergence between the coding regions of homoeologous genes in *Gossypium* typically is in the range of 1% (21, 37), and that in many cases predicted amino acid sequences are identical. This is the case for *adhA*, for example, whereas for gene B5 (an oxalate oxidase) only two amino acid differences and a two amino acid indel distinguish the two duplicates in *G. hirsutum*. Some genes may be silenced for dosage reasons (48). Alternatively, some expression variation may be functionally and selectively immaterial, reflecting instead an evolutionarily more passive side-effect of higher-order mechanistic processes that perhaps are global in scope. Future studies that include analyses of protein function and stoichiometries may shed light on these questions.

This investigation has been aided by a grant from the Jane Coffin Childs Memorial Fund for Medical Research, of which K.L.A. is a postdoctoral fellow. Financial support for this research was provided by a U.S. Department of Agriculture grant to J.F.W. and the Plant Sciences Institute of Iowa State University.

1. Ohno, S. (1970) *Evolution by Gene Duplication* (Springer, New York).
2. Gu, X., Wang, Y. & Gu, J. (2002) *Nat. Genet.* **31,** 205–209.
3. McLysaght, A., Hokamp, K. & Wolfe, K. H. (2002) *Nat. Genet.* **31,** 200–204.
4. Wolfe, K. H. & Shields, D. C. (1997) *Nature* **387,** 708–713.
5. Wendel, J. F. (2000) *Plant Mol. Biol.* **42,** 225–249.
6. Soltis, P. S. & Soltis, D. E. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 7051–7057.
7. Vision, T. J., Brown, D. G. & Tanksley, S. D. (2000) *Science* **290,** 2114–2117.
8. Leitch, I. J. & Bennett, M. D. (1997) *Trends Plant Sci.* **2,** 470–476.
9. Grant, V. (1981) *Plant Speciation* (Columbia Univ. Press, New York).
10. Stephens, S. G. (1951) *Adv. Genet.* **4,** 247–265.
11. Stebbins, G. L. (1950) *Variation and Evolution in Plants* (Columbia Univ. Press, New York).
12. Rieseberg, L. H. (2001) *Curr. Biol.* **11,** R925–R928.
13. Liu, B. & Wendel, J. F. (2002) *Curr. Genomics* **3,** 489–505.
14. Song, K., Lu, P., Tang, K. & Osborn, T. C. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 7719–7723.
15. Liu, B., Vega, J. M. & Feldman, M. (1998) *Genome* **41,** 535–542.
16. Liu, B., Vega, J. M., Segal, G., Abbo, S., Rodova, M. & Feldman, M. (1998) *Genome* **41,** 272–277.
17. Ozkan, H., Levy, A. A. & Feldman, M. (2001) *Plant Cell* **13,** 1735–1747.
18. Shaked, H., Kashkush, K., Ozkan, H., Feldman, M. & Levy, A. A. (2001) *Plant Cell* **13,** 1749–1759.
19. Wendel, J. F. & Cronn, R. C. (2003) *Adv. Agron.* **78,** 139–186.
20. Cronn, R. C., Small, R. L., Haselkorn, T. & Wendel, J. F. (2002) *Am. J. Bot.* **84,** 707–725.
21. Senchina, D., Alvarez, I., Cronn, R., Liu, B., Rong, J., Noyes, R., Paterson, A. H., Wing, R. A., Wilkins, T. A. & Wendel, J. F. (2003) *Mol. Biol. Evol.*, in press.
22. Zhao, X.-P., Si, Y., Hanson, R. E., Crane, C. F., Price, H. J., Stelly, D. M., Wendel, J. F. & Paterson, A. H. (1998) *Genome Res.* **8,** 479–492.
23. Wendel, J. F., Schnabel, A. & Seelanan, T. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 280–284.
24. Soltis, D. E. & Soltis, P. S. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 8089–8091.
25. Gastony, G. J. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 1602–1605.
26. Pichersky, E., Soltis, D. & Soltis, P. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 195–199.
27. Ford, V. & Gottleib, L. (2002) *Evolution* **56,** 699–707.
28. Pikaard, C. S. (1999) *Trends Genet.* **4,** 478–483.
29. Chen, Z. J. & Pikaard, C. S. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 3442–3447.
30. Comai, L., Tyagi, A. P., Winter, K., Holmes-Davis, R., Reynolds, S. H., Stevens, Y. & Byers, B. (2000) *Plant Cell* **12,** 1551–1567.
31. Lee, H.-S. & Chen, Z. J. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 6753–6758.
32. Kashkush, K., Feldman, M. & Levy, A. A. (2002) *Genetics* **160,** 1651–1659.
33. Force, A., Lynch, M., Pickett, F. B., Amores, A., Yan, Y.-L. & Postlethwait, J. (1999) *Genetics* **151,** 1531–1545.
34. Lynch, M. & Force, A. (2000) *Genetics* **154,** 459–473.
35. Beasley, J. O. (1940) *Am. Nat.* **74,** 285–286.
36. Wilkins, T. A. & Smart, L. B. (1996) in *A Laboratory Guide to RNA: Isolation, Analysis, and Synthesis*, ed. Krieg, P. A. (Wiley–Liss, New York), pp. 21–41.
37. Cronn, R., Small, R. L. & Wendel, J. F. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 14406–14411.
38. Small, R. L. & Wendel, J. F. (2000) *Genetics* **155,** 1913–1926.
39. Cedroni, M. L., Cronn, R. C., Adams, K. L., Wilkins, T. A. & Wendel, J. F. (2003) *Plant Mol. Biol.* **51,** 313–325.
40. Cronn, R. & Adams, K. L. (2003) *BioTechniques*, in press.
41. Liu, B., Brubaker, C. L., Mergeai, G., Cronn, R. C. & Wendel, J. F. (2001) *Genome* **44,** 321–330.
42. Madlung, A., Masuelli, R., Watson, B., Reynolds, S., Davison, J. & Comai, L. (2002) *Plant Physiol.* **129,** 733–746.
43. Richards, E. & Elgin, S. (2002) *Cell* **108,** 489–500.
44. Wolffe, A. P. & Matzke, M. A. (1999) *Science* **286,** 481–486.
45. Mittelsten Scheid, O., Jakovleva, L., Afsar, K., Maluszynska, J. & Paszkowski, J. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 7114–7119.
46. Cubas, P., Vincent, C. & Coen, E. (1999) *Nature* **401,** 157–161.
47. Lynch, M., O'Hely, M., Walsh, B. & Force, A. (2001) *Genetics* **159,** 1789–1804.
48. Osborn, T. C., Pires, J. C., Birchler, J. A., Chen, Z. J., Lee, H.-S., Comai, L., Madlung, A., Doerge, R. W., Martienssen, R. & Colot, V. (2003) *Trends Genet.*, in press.